518037

USES OF SPATIAL AUDIO IN SONAR

Prepared by
Brian E. Cooke

GENERAL DYNAMICS CANADA
3785 RICHMOND ROAD
NEPEAN, ONTARIO
K2H 5B7

Contract No.: W7711-017719/001/TOR

On behalf of
DEPARTMENT OF NATIONAL DEFENCE

Defence and Civil Institute of Environmental Medicine

Contract Report
DCIEM No. CR 2002-054
February 2002

© Her Majesty the Queen as represented by the Minister of National Defence, 2002
© Sa majesté la reine, représentée par le ministre de la Défense nationale, 2002

104-06971/

USES OF SPATIAL AUDIO IN SONAR

Prepared by
Brian E. Cooke

Author Affiliation:
GENERAL DYNAMICS CANADA
3785 RICHMOND ROAD
NEPEAN, ONTARIO
K2H 5B7
613.596.7351

Contractor/Company:
GENERAL DYNAMICS CANADA
3785 RICHMOND ROAD
NEPEAN, ONTARIO
K2H 5B7

Contract No.: W7711-017719/001/TOR

On behalf of
DEPARTMENT OF NATIONAL DEFENCE

DCIEM Scientific Authority:
G. Robert Arrabito
Defence and Civil Institute of Environmental Medicine
1133 Sheppard Avenue W
Toronto, ON, M3M 3B9
416.635.2033

The scientific or technical validity of this Contract Report is entirely the responsibility of the contractor and the contents do not necessarily have the approval or endorsement of Defence R&D Canada.

Defence and Civil Institute of Environmental Medicine

Contract Report

DCIEM No. CR 2002-054

February 2002

ABSTRACT

Underwater submarine warfare has imposed an increased threat in present-day military operations. The Canadian Department of National Defence and the defence forces in allied nations are presently investigating commercial off-the-shelf technology for enhancing the presentation of sonar data in the auditory modality in order to improve operator efficiency. One identified technology is three-dimensional (3-D) audio. A 3-D sound gives the listener the perception of the sound emanating outside of his/her head at a particular point in space even though the signal is presented over headphones. The use of 3-D audio as opposed to conventional aural presentation of sonar data is expected to result in improved operator detection and classification of enemy targets. Some recommendations and their implications are discussed in this report.

RÉSUMÉ

La guerre sous-marine mettant en jeu des sous-marins représente une plus grande menace pour les opérations militaires d'aujourd'hui. Le ministère canadien de la Défense nationale et les forces de défense des nations interalliées étudient actuellement la technologie de série servant à améliorer la présentation des données sonar dans la modalité auditive afin de rendre l'opérateur plus efficace. Une des technologies identifiées est l'audio tridimensionnel (3-D). Un signal sonore 3-D donne à la personne qui écoute l'impression d'entendre un son émanant d'une source hors de sa tête, situé à un point particulier de l'espace, bien que le signal passe par des écouteurs. On espère que l'utilisation de l'audio 3-D au lieu de la présentation sonore classique des données sonar réussira à améliorer la détection et la classification meilleures des objectifs ennemis par l'opérateur. Certaines recommandations et leurs implications font l'objet du présent rapport.

EXECUTIVE SUMMARY

There has been a renewed interest to assist the sonar operator in becoming more efficient and affective in antisubmarine warfare (ASW). This interest stems from a change in the types of threat faced by warships. The threat change is a result of two major events. First, submarine warfare has moved from a mostly deep water environment with few contacts (i.e., acoustic sources) to one that includes littoral (shallow) waters with many contacts. Second, present-day submarines are very quiet compared to the noisier nuclear powered submarines used in the Cold War era.

The change in the types of threat faced by warships has made it more difficult for the sonar operator to detect enemy targets. Although the use of audio appears to be of limited use given the advancement of digital signal processing, it still has a significant role in sonar. At present, the operator still has an advantage over computers when detecting and classifying short duration or transient auditory signals such as hull popping (e.g., caused by a submarine changing depth), 'clanking' (e.g., caused by dropping a wrench), engine start-up sequences, squeaks (e.g., caused by rudder motion), blowing ballast, explosions, and opening of torpedo doors. Such transients can greatly alert the sonar operator to the presence of an enemy target. The challenge for affective ASW is to engineer an appropriate acoustic environment so that the probability of successfully detecting the enemy is improved as much as possible. Some approaches that can assist the sonar operator are discussed in this report (Contract No. W7711-017719/001/TOR).

One present-day commercial-off-the-shelf technology that may lead to improved sonar operator efficiency is three-dimensional (3-D) audio. A 3-D sound gives the listener the perception of the sound emanating outside of his/her head at a particular point in space even though the signal is presented over headphones. Although this technology can be still considered as being in its infancy, there are potential advantages for its use that may increase sonar operator performance. For example, the use of 3-D audio compared to conventional aural presentation of sonar data may lead to quicker detection and classification of targets by increasing signal-to-noise ratios, assisting the simultaneous monitoring of multiple data channels and aiding the discrimination of the direction of the target. For these reasons, the Canadian Department of National Defence and the defence forces in allied nations are presently investigating 3-D audio technology for use in operational environments.

This report concludes by examining a number of factors that should be considered in order to improve the effectiveness of the presentation of sonar data in the auditory modality. Both factors independent and dependent of 3-D audio are detailed. Factors independent of 3-D audio include the processing of the raw data in order to ensure the frequencies are within the capability of normal hearing, and allowing the operator to replay selected segments of heard sound streams in order to reduce operator stress and fatigue. Factors dependent of 3-audio include undertaking software and hardware modifications to existing sonar sets used in the Canadian Forces in order to accommodate the presentation of sonar data in 3-D audio. Laboratory and field trials prior to deployment should be conducted to test the robustness of the proposed recommendations. Implementation of the proposed recommendations is expected to lead to increased sonar operator performance in the Canadian Forces.

SOMMAIRE

Les menaces qu'affrontent les navires de guerre ont changé, ce qui a suscité un nouvel intérêt pour aider l'opérateur sonar à devenir plus efficient et plus efficace dans la guerre anti-sous-marine (GASM). Deux événements majeurs ont transformé les menaces. Premièrement, la guerre sous-marine ne se passe plus dans un environnement d'eaux essentiellement profondes avec peu de contacts (c.-à-d. des sources acoustiques) mais se déroule dans un milieu qui englobe les eaux côtières (peu profondes) avec de nombreux contacts. Deuxièmement, les sous-marins modernes sont très silencieux par rapport aux sous-marins nucléaires utilisés pendant l'ère de la Guerre froide.

Le changement des types de menaces qu'affrontent les navires de guerre a compliqué la détection d'objectifs ennemis par l'opérateur sonar. Bien que l'utilisation d'audio semble être limitée, vu les progrès du traitement numérique des signaux, elle joue encore un rôle important dans la détection sonar. À l'heure actuelle, l'opérateur a toujours un avantage par rapport aux ordinateurs dans la détection et la classification de signaux sonores de courte durée ou transitoires tels que claquements de coque (p. ex. ceux causés par un sous-marin qui change de profondeur), bruits métalliques (p. ex. ceux causés par la chute d'une clé), séquences de démarrage de moteur, grincements (p. ex. ceux causés par le mouvement du gouvernail), délestage, explosions et ouverture des sas de torpille. Ces signaux transitoires peuvent grandement alerter l'opérateur sonar à la présence d'un objectif ennemi. Le défi à relever pour une GASM efficace consiste à créer un environnement acoustique approprié en vue d'augmenter autant que possible la probabilité de réussite dans la détection de l'ennemi. Certaines approches qui peuvent aider l'opérateur sonar sont étudiées dans le présent rapport (n° de contrat W7711-017719/001/TOR).

Une des technologies commerciales courantes qui peuvent rendre l'opérateur sonar plus efficace est l'audio tridimensionnel (3-D). Un signal sonore 3-D donne à la personne qui écoute l'impression d'entendre un son émanant d'une source hors de sa tête, situé à un point particulier de l'espace, bien que le signal passe par des écouteurs. On considère que cette technologie est encore au stade de l'enfance, mais elle présente des avantages potentiels qui peuvent améliorer la performance de l'opérateur sonar. Ainsi, l'utilisation de l'audio 3-D au lieu de la présentation sonore classique des données sonars peut aboutir à une détection et à une classification plus rapides des objectifs en augmentant les rapports signal/bruit, facilitant la surveillance simultanée de plusieurs canaux de données et aidant à distinguer la direction de l'objectif. Voilà pourquoi le ministère canadien de la Défense nationale et les forces de défense des pays alliés étudient actuellement la technologie de l'audio 3-D en vue de son utilisation dans des environnements opérationnels.

Ce rapport se termine par l'examen d'un certain nombre de facteurs dont on devrait tenir compte pour rendre plus efficace la présentation des données sonars dans la modalité auditive. Y sont décrits les facteurs indépendants aussi bien que les facteurs dépendants de l'audio 3-D. Les facteurs indépendants de l'audio 3-D comprennent le traitement des données brutes, pour garantir que les fréquences se trouvent dans les limites de l'ouïe normale, et la relecture par l'opérateur de segments sélectionnés des flux sonores écoutés, pour réduire le stress et la fatigue de l'opérateur. Les facteurs dépendants de l'audio 3-D comprennent la réalisation de modifications logicielles et matérielles des dispositifs sonars existants utilisés par les Forces canadiennes, afin de permettre la présentation de données sonar en audio 3-D. Il faudrait effectuer des essais en laboratoire et sur le terrain avant le déploiement pour tester la consistance

des recommandations proposées. On s'attend à ce que la mise en œuvre des recommandations proposées donne lieu à une performance améliorée des opérateurs sonars des Forces canadiennes.

TABLE OF CONTENTS

Para		Page
	ABSTRACT / RÉSUMÉ	i
	EXECUTIVE SUMMARY	ii
	SOMMAIRE	iii
1.	**SCOPE**	**1**
1.1	*Identification*	*1*
1.2	*Rationale for and Introduction of this Report*	*1*
1.2.1	*Definition of Spatial Audio or 3D Audio*	*2*
1.3	*Document Layout*	*3*
2.	**HISTORICAL OVERVIEW**	**4**
2.1	*Introduction*	*4*
2.1.1	*Passive Sonar – Terms of Reference*	*4*
2.1.2	*Active Sonar – Terms of Reference*	*5*
2.2	*Audio Issues Overcome by Visual Presentations*	*6*
2.2.1	*Passive Sonars*	*6*
2.2.2	*Active Sonars*	*7*
2.2.3	*Common Between Active and Passive Sonars*	*7*
2.3	*Issues as Candidates for Audio*	*8*
2.3.1	*Operator as a Part of the Team*	*8*
2.3.1.1	*Aircraft Background Noise Levels*	*9*
2.3.2	*Listening to Sonar Audio*	*9*
2.3.2.1	*Aircraft Sonar Sets*	*10*
2.3.2.2	*Ship's Sonar Sets*	*11*
2.3.3	*Transient Detections*	*11*
2.3.4	*Expected Benefits of Spatial Audio to the Acoustic Operator*	*12*
3.	**PRESENT STATE OF THE ART**	**13**
3.1	*Sonobuoy Processor Highlevel Block Diagram*	*13*
3.1.1	*Hardware Limitations*	*14*
3.2	*Towed Array Passive Sonar*	*15*
3.2.1	*Common Audio Elements for Ship Sonars*	*16*
3.2.2	*Hardware Limitations*	*17*
3.3	*Hull Mounted Active Sonar*	*18*
3.3.1	*Hardware Limitations*	*19*
4.	**RECOMMENDATIONS**	**20**
4.1	*Non-Spatial Audio Factors*	*20*
4.1.1	*Improvements to Early Stages of Audio Processing*	*20*
4.1.2	*Operator Fatigue and Historical Data Replay*	*21*
4.2	*Spatial Audio Benefits*	*21*
4.3	*Operator or Human Factors*	*22*

PROPRIETARY
General Dynamics Canada
"Use, duplication or disclosure of this document or any of the information contained herein is subject to the restriction on the title page of this document."

4.3.1 General HRTF or Personal HRTF 22

TABLE OF CONTENTS (CONT'D)

4.3.2 Number of Sound Sources 23
4.3.3 Operator Comfort 24
4.3.4 Merging of Voice and Data 24

Para Page

4.3.5 Audio Control interface 24
4.4 Scope of Change 25

5. CONCLUSION 26

6. APPLICABLE DOCUMENTS 28
6.1 Canadian Government Documents 28
6.2 General Dynamics Canada Documents delivered to the Crown 28
6.3 Commercial Publications 29

LIST OF ILLUSTRATIONS

Figure Page

1 A Sonobuoy Active/Passive Sonar Block Diagram 14
2 CANTASS a Passive Sonar Block Diagram 16
3 Block Diagram of the Audio CCA used in CANTASS and AN/SQS-510 18
4 AN/SQS-510 an Active Sonar Block Diagram 19

PROPRIETARY
General Dynamics Canada
"Use, duplication or disclosure of this document or any of the information contained herein is subject to the restriction on the title page of this document."

1. SCOPE

1.1 Identification

This report reviews the use of audio to present information to a sonar (acoustic) operator. The acoustic operator needs to manage voice communication, alerts, and sonar derived audio streams. The use of spatial audio, see definition below, is explored as a solution to operator loading issues experienced with the current approaches. The report has been prepared under funding from the Defence and Civil Institute of Environmental Medicine (DCIEM) – Contract W7711-017719/001/TOR.

1.2 Rationale for and Introduction of this Report

The acoustic operator performs an important role in the operation of deployed sonars. This role is assisted by the use of audio to help manage the interpretation of sonar information. Several competing types of information are presented to the operator in audio format – voice, alerts, and sonar streams. Historically the role of audio was very important, but many of the roles associated with audio have been replaced by other techniques. This review has been triggered as a result of two significant shifts – a drive for greater operator efficiency and the reduced cost of technology:

a. A change in the types of threats being faced by warships and aircraft has increased the interest in optimising the use of audio. The threat change is a result of moving from a mostly deep-water environment with few contacts to one that includes shallow water (Littoral) with many contacts. Furthermore, the increase in importance of dealing with very quiet submarines compared to the noisier nuclear powered submarines of the Cold War era has increased the load on the operator. A related factor is the need to reduce the number of personnel required to operate sonar. This reduction requires that an operator become 'more efficient'.

b. Improvements in the audio stream are anticipated to help address operator loading. Innovative approaches are now available, including spatial audio. A significant cost reduction in spatial audio hardware and software development environments has made the deployment of spatial audio cost effective. The commercial market place, specifically the home PC Market, has driven this cost reduction. The main use of the PC hardware is computer-based games, for example, see http://www.gamasutra.com. This highly competitive market has driven the cost of audio processors from significant levels down to the point that a home computer user would consider a purchase. Further evidence of the commercial acceptance of spatial audio is that support for spatial audio is available[1] for the JAVA programming language.

[1] see http://java.sun.com/products/java-media/3D/forDevelopers/J3D_1_2_API/j3dapi/javax/media/j3d/package-summary.html

PROPRIETARY
General Dynamics Canada
"Use, duplication or disclosure of this document or any of the information contained herein is subject to the restriction on the title page of this document."

1.2.1 Definition of Spatial Audio or 3D Audio

Spatial Audio, also known as 3D Audio, uses signal processing techniques to introduce auditory localisation cues into the sound stream. The aim of this processing is to 'place' a specific sound at a specific location in three-dimensional virtual space around the listener. Different processing is required if speakers or headsets are used. Audio appropriately synchronised with video is an important component of a virtual reality experience. Even with only left and right audio signals being presented with a headset to a listener, the sound can appear to come from the front, from behind, up high or down low. The quality of this subjective positioning depends on the matching of the Head Related Transfer Function[2] (HRTF), reference 6.3e, to the actual listener. Eight localisation cues help the listener locate the position of a sound source in space (from http://www.hitl.washington.edu/scivw/EVE/index.html).

a. **Pinna response:** the effect that the external ear, or pinna, has on incoming sound. Typically captured as part of the HRTF.

b. **Interaural time difference:** the time delay between sounds arriving at the left and right ears.

c. **Head shadow:** the modification to sound that passes through or around the head in order to reach an ear.

d. **Shoulder echo:** sound in the range of 1-3 kHz is reflected from the upper torso of the human body.

e. **Head motion:** movement of the head and corresponding changes in the sound helps determine the direction of a sound source.

f. **Early echo response:** occurs in the first 50-100 ms of a reception.

g. **Reverberation:** reflections from surfaces around the listener.

h. **Vision:** helps establish possible locations and possibly confirms the location or direction of a sound.

---

2 The HRTF is a fundamental component of spatial audio. The HRTF captures the modifications to sound caused by the presence of an individual prior to reception by the inner ear. The modifications depend on many factors – position (bearing, elevation, and distance), and signal frequency. The modifications to the sound create position cues that the brain can interpret. The brain uses these cues to create the perceived positions of these sounds.

PROPRIETARY
General Dynamics Canada
"Use, duplication or disclosure of this document or any of the information contained herein is subject to the restriction on the title page of this document."

1.3 Document Layout

The format of this document is as follows:

a. Section 1 – An **introduction** to the purpose of this report and the document layout.

b. Section 2 – **Historical Overview**. This section contains a brief overview of the methods used to process and present active and passive sonar data. A discussion of the apparent preference for the use of a visual modality is given.

c. Section 3 – **Present State of the Art**. An analysis of the sonars deployed on Canadian Ships is given. A variety of factors that are associated with sonar audio will be discussed – hardware limitations, operational needs, etc.

d. Section 4 – **Recommendations**. An analysis of the cost and feasibility of improving the audio presentation of sonar data is provided. Human factors, scope of change, and differences between active and passive sonar needs are described.

e. Section 5 – **Conclusions**.

f. Section 6 – A list of documents **referenced** in this report.

PROPRIETARY
General Dynamics Canada
"Use, duplication or disclosure of this document or any of the information contained herein is subject to the restriction on the title page of this document."

2. HISTORICAL OVERVIEW

2.1 Introduction

The detection of underwater sound with the human ear has been documented for centuries. A good review of undersea acoustics up to 1950 is available in reference 6.3d. In *Principals of Underwater Sound*, reference 6.3a, the following comment is found:

> *"One of the earliest references to the fact that sound exists beneath the surface of the sea as well as the air above, occurs in one of the notebooks of ... Leonardo da Vinci. In 1490 ... he wrote: If you cause your ship to stop, and place the head of a long tube in the water and place the outer extremity to your ear, you will hear ships at a great distance from you"*

2.1.1 Passive Sonar – Terms of Reference

The quote above introduces a passive sonar, one that only listens to the natural emissions from a platform of interest. The evolution of passive sonar over the years has improved performance by exploiting specific characteristics of the sound emitted from platforms of interest. During the Cold War, significant advances in digital signal processing used the Fourier Transform to produce detectors that excel at detecting stationary narrowband lines in a spectrum. These emissions were typical of the mechanical equipment found on Soviet submarines. The stationary nature of the lines allowed the signal to be averaged (or integrated[3]) over 'seconds' to extract the signal from the noise. Thus detection at long ranges was possible. In almost all cases the results of the signal processing were displayed on a video screen or printed on paper charts. The operator's eye was used to perform a final level of integration and the operator typically made the detection by looking at a LOFARgram[4]. Typical frequencies used for passive ship detection fall between 50 – 500 Hz (reference 6.3a, page 207). Audio is still used to augment the visual presentation in particular, by detecting transients and helping classifying the emissions.

The first example of a deployed passive sonar in the Canadian Navy is the AN/SQR-501 CANTASS, reference 6.2a. The Canadian Towed Array Sonar System (CANTASS) consists of a passive towed array, (towed well behind a ship), the handling gear, and the processing and operator interface (visual display, inputs devices and audio signals). CANTASS is deployed on all twelve of the Canadian Navy Halifax Class ships. CANTASS receives passive emissions from platforms and provides the operator with the ability to detect, classify and track the platform. Like all passive sonars, CANTASS provides signal characteristics and signal bearing.

---

3 The actual integration period used in military passive sonars is usually classified as a CONFIDENTIAL parameter. The frequency span and processing frequency cell width are also classified.

4 A LOFARgram is a plot with frequency across the horizontal axis and time on the vertical axis. The data are plotted in rows of frequency cells, one row for each integration period. The intensity of a frequency cell in the LOFARgram represents the amount of energy detected at that frequency cell and that integration period. With such a plot, a continuous emission at a constant frequency appears as a vertical line.

PROPRIETARY
General Dynamics Canada
"Use, duplication or disclosure of this document or any of the information contained herein is subject to the restriction on the title page of this document"

The range to the platform is not measured directly by the sonar. Techniques are available that can estimate the range, but these usually require a series of measurements over time. In another example, 27 Modular VME Acoustic Signal Processors (MVASP) for the CP-140 Aurora Maritime Patrol Aircraft fleet have been recently placed on order. These systems use sonobuoys (often shortened to just "buoys") as the sensors in the water. Both active and passive buoys are supported. The data are radioed to the aircraft for processing, display and interpretation. In passive mode, the location of the platform can be deduced by which buoys, of a buoy field, are detecting the signals of interest. Normally a number of buoys are deployed. A final example is the ~20 AN/UYS-503 processors carried on the Canada's fleet of Sea King Helicopters. Like the MVASP, these processors also receive signals from the sonobuoys that have been dropped in the water. General Dynamics Canada Ltd has manufactured all of these passive sonar processors.

2.1.2 Active Sonar – Terms of Reference

Active sonar deliberately creates an acoustic pulse that is propagated by the water around the sonar. Then the sonar listens for an echo. The time between the creation of the acoustic pulse and the reception of the echo is proportional to the range of the target. The bearing of the received echo provides the bearing of the target. Thus active sonar provides both range and bearing of a target. Like passive sonar, digital signal processing techniques have improved the effective ranges of active sonar. The generation of the transmitted waveform (or ping) is typically performed with digital control. The precise specification of the ping – its starting frequency, frequency sweep rate, ping length in time, and power level, allows a matched filter[5] processor to be used to extract echoes from the target of interest. This type of processing is more effective at identifying weak echoes than the 'energy detector' of the human ear. For example, in the AN/SQS-510 sonar, reference 6.2b, the typical frequency used has pings centred at 7.2 kHz with a 300 Hz sweep and 320 ms ping duration. The type of ping sweep used is a linear period sweep – the instantaneous period varies as a linear function of time during the ping.

Previous versions of active sonar operated by the Canadian Navy, e.g., AN/SQS-505, relied on the audio channel to assist the operator in understanding the acoustic scene. The AN/SQS-505 used much shorter pings, 10 - 40 ms, and a fixed frequency tone (CW). The operator was able to detect echo frequency shifts (caused by the Doppler Effect) to deduce the rate of closing or opening of the target. The longer, 400 ms, CW ping of the more modern AN/SQS-510 sonar presents a bigger challenge to the operator. This observation has been explained by the psycho-acoustic properties of human hearing – sounds of duration of less than ~100 ms are interpreted differently than sounds of longer duration[6]. Increased reverberation, a characteristic of littoral water, challenges the operator even more and may shift the 'best' ping length for audio.

---

5 from Urick page 384, reference 6.3a: "... the optimum receiver for white noise is the *matched filter* whose impulse response is the same as the waveform of the known signal reversed in time. A practical realization of the matched filter is the crosscorrelator. ...in which the signal plus noise is correlated, or multiplied, with a noise-free replica of the known signal."

6 Sharon McFadden, DCIEM (2000) – private communication.

PROPRIETARY
General Dynamics Canada
"Use, duplication or disclosure of this document or any of the information contained herein is subject to the restriction on the title page of this document."

The use of frequency sweeps during FM pings, like those of the AN/SQS-510, resulted in reduced use of audio because the ping was long and the ping structure too complex for simple interpretation by a human, see the Audio Clip in section 2.2.2. It is of some interest that other mammals, bats and whales, are able to 'process' FM pings with great success, reference 6.3c. FM pings have greater importance in shallow water operations because they can reduce the interference from reverberation much better than with CW pings.

The processing and display components of the AN/SQS-510 are produced by General Dynamics Canada. This sonar uses a four foot diameter by 4 foot high transducer that is either mounted below the keel of a Frigate sized ship, the Hull Mounted Sonar (HMS) version, or towed behind a ship in the Variable Depth Sonar (VDS) version. The Canadian Navy's Halifax Class ships have the HMS version and the Canadian Navy's Iroquois Class ships have both the HMS and VDS versions. The two aircraft systems, MVASP and the AN/UYS-503 can also process active sonobuoys. These allow ranges to be determined more directly than in passive only operations.

2.2 Audio Issues Overcome by Visual Presentations

Although the most natural method for detecting either passive emissions or active echoes from a target is with the human ear, a variety of factors have encouraged the use of alternative schemes:

2.2.1 Passive Sonars

a. **Operator Fatigue.** The detection of weak tonal lines in a background of noise requires significant concentration by an operator. Unlike listening to music or voice, the sound is basically noise – this noise is tiring. The operator's success depends on training and ability to concentrate. Pressures to reduce staffing levels leads to longer times with a given operator seated at the sonar console. In contrast, a computer can perform the digital signal processing repeatedly and reliably. The various video displays available to the operator, reference 6.1a, still require a level of concentration, but are tolerant of short lapses in attention.

b. **Passive detection integration period.** The human ear has a natural integration period that appears to lie between 0.1 and 1 second, reference 6.3a – section 12.10. But longer integration periods are often required to recover sinusoidal signals buried in broadband noise. Digital signal processing allows an arbitrarily long integration period to be created. Visual presentations, historically on paper, but now on either CRT or LCD monitors[7] are used to present the results to the operator.

---

7 Historically bulky evacuated glass CRT Monitors have been used to create the visual display. This was the only technology available. Over the last few years, the LCD Monitor has become available and in recent years has become much more affordable. The advantages of LCD monitors are: light weight, lower electrical demands, and smaller footprint. At the time of writing this report, commercial quality LCD Monitors are still about 5 times as expensive as CRTs but the price continues to drop. LCD monitors have only recently achieved the resolution typically required for passive sonar operations (1600 x 1200) pixels. Most LCD monitors exhibit a flicker when LOFARgram type displays scroll, reference 6.3h. The limited viewing angle of some LCD monitors has been of concern historically, though some suppliers have made significant advances. The trend in sonar is definitely towards LCD monitors as the price, flicker, and viewing angle issues are resolved. Other flat panel monitor technologies, e.g., organic light emitting diodes or Field Emission Displays (FEDs), could overtake LCD technology, but at this time, none are production.

PROPRIETARY
General Dynamics Canada
"Use, duplication or disclosure of this document or any of the information contained herein is subject to the restriction on the title page of this document."

c. **Longer Memory.** The human ear has limited ability to memorise and compare signals over a long period of time. Extracting very weak signals from a noisy background can often be performed in modern sonars if sufficient storage and display is provided to support further 'signal averaging' of the data. The LOFARgram display format is the result of years of refinement and excels at presenting this historical record.

2.2.2 *Active Sonars*

a. **Interpretation of Complex Pings.** The human operator has significant difficulty extracting and understanding even the simplest frequency swept pings. These types of pings are much more important in shallow water where reverberation is typically higher than in deep water. The digital processing with a matched filter combined with a video display leads to a more effective sonar. For example, the AN/SQS-510 FM linear period sweep pings are so difficult to interpret that essentially no AN/SQS-510 operators use their headsets to listen to the pings while operating in FM.

Audio clip: A
Linear Period Modulation FM ping and
echo (5 seconds after ping) with very good signal to noise

b. **Ping History.** Although different from the passive case in detail, the human ear-memory combination cannot reliably maintain sufficient memory of recent events to support identification of weak intermittent echoes in a complex scene. The situation becomes even more difficult in Littoral waters - a feature rich environment. There are just too many echoes to track. Various video display formats have been developed to overcome this issue, reference 6.1b.

2.2.3 *Common Between Active and Passive Sonars*

a. **Audio Alerts and Voice Channel.** Operators are already using the audio channel to support voice alerts as well as voice communication with other members of the team. The dedication of audio to signal detection would lead to a type of isolation for the operators. This observation requires the *status quo* to be updated.

b. **In aircraft operations, the noise level caused by propellers and other machinery reduces the effectiveness of the operator.** The background sound leaks through the headsets to the point that it very difficult to make early detection of weak signals.

c. **Overall Operator Fatigue.** On board fixed wing aircraft, 'on position' times are typically between 4 – 8 hours with no relief rotation. Helicopter missions are shorter, up to 2 – 3 hours. On ship deployed sonars the operators are typically rotated away from the console every 20 minutes and are assigned other tasks to relieve fatigue. Then after 20 minutes they return. Typical ship watches vary - but 8 to 12 hours are possible. The ship sonars are manned 24 hours a day if appropriate.

PROPRIETARY
General Dynamics Canada
"Use, duplication or disclosure of this document or any of the information contained herein is subject to the restriction on the title page of this document."

2.3 *Issues as Candidates for Audio*

Although, the use of audio appears to be of limited use given the factors raised in the previous section, audio still has a significant role in sonar. At this time, the operator still has the advantage over computers when detecting and classifying the following short duration or transient signals, see section 2.3.3:

a. hull popping, caused by a submarine changing depth;

b. 'clanking' for example caused by dropping a wrench;

c. engine start-up sequences;

d. squeaks, for example caused by rudder motion;

e. shaft rub caused by a noisy shaft housing;

f. blowing ballast;

g. explosions; and

h. opening of torpedo doors.

The challenge is to engineer an appropriate acoustic environment so that the probability of success is improved as much as possible. Another factor to consider is that many sonar operators have developed hearing loss from exposure to high ambient noise in aircraft. The exposure to this noise happens when they slip the headset off to communicate with another operator. Some of the issues can be overcome with the use of spatial audio to separate:

a. voice channels for different individuals so that the importance of the information can be sorted quickly;

b. acoustic data streams including replay of sound clips; and

c. surveillance monitoring of a single listen direction to a full 360 degree coverage.

The following discussion elaborates on approaches that can assist the sonar operator by presenting information in a natural way. In reality, most if not all of the audio streams will likely be presented to the operator at the same time. The challenge is to identify a method of presentation that minimises the operator loading required to deal with the information.

2.3.1 *Operator as a Part of the Team*

The sonar operator is located in the Operations Room on board ship or on board an aircraft (either fixed wing or rotary wing). They are part of a team that detects and ultimately deals with targets. Part of the team's success depends on the ability to communicate effectively with each

PROPRIETARY
General Dynamics Canada
"Use, duplication or disclosure of this document or any of the information contained herein is subject to the restriction on the title page of this document."

other. Typically, detailed sonar data are exchanged electronically between members of the teams. The information is passed up through a chain of individuals and eventually is combined with other sensor data. The electronic transfer is enhanced by the ability of the operator to participate in a voice network of several sensor operators. For example, discussions between the passive sonar operator (CANTASS) and the active sonar operator (AN/SQS-510) can lead to better understanding of target characteristics.

On board aircraft sonar operators communicate with other members of the aircrew. This exchange of information is important for effective operations. The aircraft based operators have potentially longer watches and aircraft noise to deal with. In either the ship or aircraft scenario, the acoustic operators wear microphone-equipped headsets.

In many cases, the sonar operators, either on ships or aircraft, communicate by slipping off one of the earphones and use voice – bypassing the electronics. This is done to reduce the electronic voice channel traffic. This approach increases the noise level that the operator must tolerate, but has been found to be operationally effective. Any solution proposed as a result of this work needs to address the underlying causes of this procedure.

2.3.1.1 Aircraft Background Noise Levels

The airborne acoustic operator is exposed to background noise that reduces the effectiveness of any information transferred over the audio channel. Active noise cancellation technology has demonstrated dramatic improvements. Although not part of spatial audio, noise cancellation is likely a requirement for successful spatial audio capability on aircraft. This technology has advanced significantly in recent years. Commercial versions are now available, for example see:

http://www.bose.com/noise_reduction/personal/qc_headset/

for a description of a headset that is sold for US $300.

2.3.2 Listening to Sonar Audio

The operators monitor the audio channels in specific operational scenarios. Even a single audio data stream is very difficult to absorb for long periods of time yet a significant number of audio streams require monitoring. For example in an airborne environment as many as 32 sonobouys could be of interest.

The AN/SQS-510 and CANTASS sonars do not supply the sonar operator with raw data streams in audio format. The data are usually beamformed, e.g., spatial filtered, to increase signal to noise[8] and frequency shifted (heterodyned) to move the signal to a band where the ear is

[8] Sonars often operate with signal to noise ratio levels of less than 0 dB at the raw data stage. The filtering stages, either spatial filters or frequency filters, are applied to improve the signal to noise to be at least 3 dB. As an example of spatial filtering (or beamforming): a simple approximation for the array gain for a line array that is 10 wave lengths long is about 13 dB. see page 43, reference 6.3a.

PROPRIETARY
General Dynamics Canada
"Use, duplication or disclosure of this document or any of the information contained herein is subject to the restriction on the title page of this document."

sensitive[9]. In many cases passive sonars can benefit from shifting the frequencies up, while active sonars usually benefit from shifting frequencies down. In military sonars the active echoes are often frequency shifted to a band centred near 1 kHz. In active sonars, a varying gain control is applied to the data to greatly reduce the rather large dynamic range of signal amplitude following a ping. (The AN/SQS-510 front end has a 140 dB dynamic range). In passive sonar, a nearly constant gain control is used.

In airborne sonars, processed as well as raw audio data are provided to the operator. Processed audio can be beamformed to increase the signal to noise ratio. With active DICASS sonobuoys, the signal band of interest is shifted in the sonobouy down to a band centred at 800 Hz. As in the shipborne environment, continuous monitor of even one sonobuoy is very difficult for long periods of time; continuous monitoring of two sonobuoys (using left and right channels), i.e., a dichotic presentation, is even more difficult; and continuous monitoring of all sonobuoys is impossible. Consequently techniques have been developed for detecting transient signals with the purpose of alerting the operator.

More recent sonar configurations allow the operator to replay recently acquired audio clips (for example the Swedish Navy's Hydra Sonar and the MVASP sonars). Typically the last few minutes of audio data are kept for replay. This contrasts with significantly longer visual histories that are maintained. The cost of storing significantly greater audio history limits the amount stored. The automatic transient detectors now being built into sonars would be used to alert to the operator. Once alerted by the sonar, the acoustic operator can replay the sound byte of interest to confirm the detection was appropriate. The speed of replay can be real time, slower than real time or faster than real time.

The replay capability assists the operator by helping answer the question – "What was that sound?" The replay also allows the operator to use the automatic processors to scan the data for events of interest that are proposed to the operator. This reduces the fatigue associated with the operator being exposed to the mostly noise and possibly weak signals picked up by the sonar.

2.3.2.1 Aircraft Sonar Sets

The physical position of sonobuoys within a field depends on the type of pattern required to satisfy the tactics in play at the time. Thus the relative positions of the various buoys are not predetermined. Currently the operator manually steps through the signals from the deployed buoys and examines the data from each – potentially even listening to the audio from each buoy. Current practice reports that the audio from active buoys is monitored more than passive buoys. A close-in contact can be monitored in passive but with today's quiet submarines the amount of time that a passive audio contact would be held is of the order of minutes. A convenient mechanism that would allow the operator to monitor several buoys at once is of interest.

---

9 Passive sonars typically cover frequencies below 10 kHz. The attempt to get as low as possible consistent with the detection array dimensions (the exact frequencies are confidential). Lower frequencies require large detector sizes if beamforming is to be used. Active sonars often transmit at higher frequencies: long range sonars about 1-3 kHz, medium frequency sonars transmit about 10 kHz, mine hunting operate at greater than 90 kHz. Some specialised sonars operate at 500 kHz or more. Higher frequencies suffer from much greater absorption in seawater, but can provide very good spatial resolution for a given transducer size because of the short wave length.

PROPRIETARY
General Dynamics Canada
"Use, duplication or disclosure of this document or any of the information contained herein is subject to the restriction on the title page of this document."

2.3.2.2 *Ship's Sonar Sets*

The AN/SQS-510 transducer has a fixed array of sensors. The towed array, CANTASS, has sensors fixed in the array, but they can snake through the water or form a curve if the ship alters course or speed. Most of the outputs of the sonars are arranged in beams of data. Audio presented to the operator is usually beamformed audio. (In most cases, audio from an individual hydrophone channel is only used by a maintainer trying to locate a defect.) The audio can be generated for an arbitrary beam. In principal audio can be generated for all beams at once – but this is not done in most deployed systems. The beams are numbered and follow a simple pattern. On the HMS active sonar, the 36 beams are spaced every 10 degrees in azimuth around the ship. For the towed array 43 beams are distributed with look directions of between dead ahead and dead astern, but they are not equally distributed because of the consequences of a linear array. There is a higher density of beams near broadside. The large number of beams overwhelms the capability of an OMNI-audio channel.

2.3.3 *Transient Detections*

Submarine builders have and continue to develop techniques to reduce the effectiveness of sonars, either passive (e.g., careful mounting of equipment) or active (anti-reflection tiles). As a result the signal intensity received by the sonar has become less. One type of emission, transient emission, has remained as a problem for both the submarine builder and the sonar operator. The transient emission is difficult to detect because of the variety of possible signal types. The creation of a matched filter, tuned to all transients, has not and likely never will be found. There is one solution that is available and usually out performs any machine detection – the human ear. Human hearing is well tuned to detecting and classifying transients. The problem in sonar is to present the data in a way that is not overwhelming.

Transients are a class of emissions that are characterised with durations of 'seconds or less'. This includes a very broad range of signals – hence the difficulty in designing a universal detector. Most transients have energy spread over a broad range of frequencies. Examples of transients of interest include: opening of torpedo doors, hull noises caused by depth changes, and other intermittent machinery noises. Most of the actual sounds are classified. A suitable typical sound has been found on a web site:

http://www.geocities.com/CapeCanaveral/1046/audio1.html

On this web site, a sound file has been located that captures a sequence of transients made by biologicals – Carpenter Fish. The sounds have a higher density of events than typically found, but are still felt to be representative. The sound clip is included below:

carp3.wav

A sound clip that captures the emissions from a school of Carpenter Fish.

As stated before, transients are difficult to detect. If signals are integrated in an attempt to enhance signal to noise ratios, the transients are usually 'integrated out' because the integration periods are usually longer than a transient. The optimal integration period for a transient is the duration of the transient – something that is not known in advance. The human ear has an ability

PROPRIETARY
General Dynamics Canada
"Use, duplication or disclosure of this document or any of the information contained herein is subject to the restriction on the title page of this document."

to detect a wide variety of transients. Audio monitoring of a beam that is pointed in the correct direction could be successful. The human ear can be monitoring while other activities are being pursued. This alerting mechanism is a very natural way of using sound.

2.3.4 Expected Benefits of Spatial Audio to the Acoustic Operator

In the most abstract view, the benefits of spatial audio come when the situation requires the presentation of many sound streams to the operator's two ears. What has been done in the past is to just add the signals together in a mixer, OMNI-Audio, and let the operator try to figure out the result – a situation that just does not work very well. If spatial audio is used, then the signals are multiplexed into the audio stream in a way the human brain understands and can decode. The operator has a natural way of demultiplexing the data streams back to individual streams. The recovery of the different data streams supports the following benefits:

a. Increased ability to monitor several different acoustic stream of sonar data with much higher signal to noise than if an OMNI-audio approach was used;

b. Better separation of voice and acoustic streams. This is superior to piping voice in one ear and sonar acoustic data into the other. Also several different channels can be placed around this audio space – for example: a Active sonar operator to the right, passive sonar operator to the left and supervision above. This leads to a more natural way of interpreting information; and

c. The transient detection requires maximising the signal to noise ratio (SNR). But if several channels are to be monitored the traditional method sacrifices the SNR to provide coverage over many channels. The spatial audio allows the monitoring of a number of channels without completely removing the beamforming gain in SNR. This coupled with the observation that the human ear is very effective at detecting and classifying transients improves the likelihood of early correct assessment of the signal. This is expected to lead to earlier detections based on transients.

PROPRIETARY
General Dynamics Canada
"Use, duplication or disclosure of this document or any of the information contained herein is subject to the restriction on the title page of this document."

3. PRESENT STATE OF THE ART

A high level block diagram will be used to describe three different sonar examples: a sonobuoy processor, a towed array sonar and an active hull mounted sonar. All of the sonar installations currently do not have any head tracking capability – yet research has shown significant reduction in front/back confusion by the use of real time head tracking[10], reference 6.3f. The importance of head tracking for sonar applications is not known at this time.

3.1 Sonobuoy Processor Highlevel Block Diagram

The block diagram of a sonobuoy processor is provided on Figure 1. The signals from the sonobuoy are sent back to the aircraft over a RF link. The sonobuoy receiver provides a separate data stream for each sonobuoy within the range of the RF link. The data from each sonobuoy is processed in parallel.

The signal (raw data) from an individual sonobuoy is made available to the operator in audio format. If appropriate the signal can be band shifted or heterodyned to a frequency band typically centred near 800 Hz before presenting to the operator.

The signals from a single sonobuoy are demultiplexed to recover the directional response. The data are fed to an audio storage that retains the last few minutes of audio. The data from this audio storage can be replayed to the operator at 1/8, ¼, ½, 1, 2, 4, 8 times real time. In other words the information is available with time compressed or time expanded.

The demultiplexed data are also sent for beamforming. Beamforming improves the signal to noise ratio. This allows weak signals to be boosted relative to the noise by preferentially 'listening' in the desired direction. Detailed spectral analysis, extensive video history and display generation complete the processing.

Either automatic or manual Audio gain controls boost the signals to an appropriate level prior to driving the headsets. Other controls are provided (not shown on diagram) that allow the operator to select the buoy being monitored. A single sonobuoy or two sonobuoys may be followed at once. These are presented in either a dichotic[11] (binaural) or diotic[12] (mono) format. Typically the operators select 'SonoOnly' (thereby disconnecting other audio streams) in order to focus on the audio from sonobuoys.

---

10 Head Tracking requires equipment that determines the orientation of the head relative to some fixed points. The stimulus, in this case sound, would be altered by the audio system to take account of the new head orientation. In many cases test subjects have confused or interchanged the position of virtual sounds that were 'placed' behind the subject with sounds 'placed' in front of the subject. Head tracking with the associated audio corrections breaks the symmetry enough that many of the confusions are resolved.

11 dichotic - a different signal is sent to each ear. Commonly referred to as stereo.

12 diotic - affecting both ears, the same signal is presented to both ears. Commonly referred to as 'mono' though strictly speaking 'mono' means the presentation of a signal to only one ear and nothing to the other ear.

PROPRIETARY
General Dynamics Canada
"Use, duplication or disclosure of this document or any of the information contained herein is subject to the restriction on the title page of this document."

3.1.1 Hardware Limitations

The aircraft crew routinely wear headphones to reduce noise levels and allow communication. Thus wearing headphones while operating the sonar is already part of the routine. The frequency response of the headphones in use is of some concern[13]. One of the major concerns is the computing resources required to general the spatial audio signals. If more than one audio channel is required to be presented, then the processing increases even further. Currently this processing would consume computation bandwidth normally allocated to the basic processing needs. The hardware is not currently configured to deliver dichotic signals.

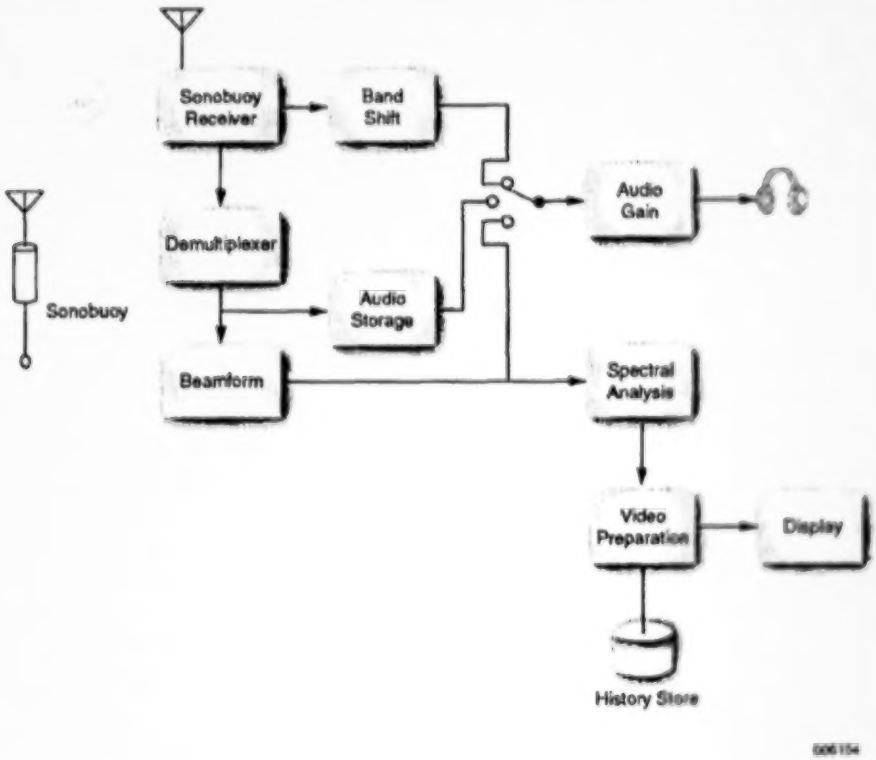


Figure 1. A Sonobuoy Active/Passive Sonar Block Diagram

Depending on the approach followed, processing resources could be diverted to audio or additional processing provided by adding more processing elements.

---

13 The headphones have an upper cut-off frequency near 4.5 kHz. Above this frequency, the response falls off quickly. The existing communications channels may also present a challenge. They have an upper cut-off near 3.5 kHz. These two limitations can contribute to poor spatial audio.

PROPRIETARY
General Dynamics Canada
"Use, duplication or disclosure of this document or any of the information contained herein is subject to the restriction on the title page of this document."

Another issue – not specifically hardware related, is the need to control the audio sources. Multiple audio sources would need to be identified and distributed in the virtual space about the operator. The mechanism to control this would require some design to create a suitable interface that takes account of the various factors that face crew members in an aircraft. This interface design is further complicated by the fact that the buoys are distributed in a variety of different buoy field geometric layouts. The operator needs to be able to track them in the virtual audio space as well as in the geographic space.

3.2 Towed Array Passive Sonar

The block diagram of a towed array passive is shown in Figure 2. The digitised signals from the towed array hydrophone channels are sent up the tow cable to the array receiver. The array receiver can generate audio for a selected hydrophone channel. The sonar maintainer can listen to this individual hydrophone channel for fault finding, but this is not used during normal operations. (Faults in the array often are characterised by either a dead channel or an extremely noisy or howling channel. A maintainer can identify the bad channel quickly by listening to each channel in succession and noting channels that are significantly different from neighbours.) The acoustic data from the array receiver are beamformed into 43 non-overlapping beams. The two CANTASS operators sit at two dual screen displays and are provided with two audio channels that can be assigned to operator selected bearings. The operator then can choose to listen to either channel, but not both at the same time. The audio headsets have the left and right channels wired together providing a diotic presentation, i.e., not currently configured to deliver dichotic signals.

The operators use the video presentation to monitor the data for narrowband lines and broadband emissions from targets of interest. Historically the narrowband lines were of greatest interest, but modern targets, submarines, are so quiet that broadband and transients are often the only methods available for detection. A video history allows the operator to go back in time (a classified amount of time) to re-examine signals of interest. The two operators are required to process the data on all the beams. They often search by stepping through the beams one at a time. The fatigue on the operator is a concern. In many situations the operators are relieved often and given some other duty as a temporary break from the intense concentration required.

PROPRIETARY
General Dynamics Canada
"Use, duplication or disclosure of this document or any of the information contained herein is subject to the restriction on the title page of this document."

3.2.1 *Common Audio Elements for Ship Sonars*

The audio interface hardware in the deployed AN/SQS-510 and CANTASS sonars is identical, see Figure 3.

The audio circuitry consists of 4 independent, 8 bit channels. The data used for the audio come from a so-called Mass Memory store, where the four channels are packed into the same 32 bit word. The audio interface is capable of generating audio output over the frequency range from near DC to 8.6 kHz. The digital data are read from Mass Memory by the processor and written into the audio First In First Out (FIFO) buffer. This allows the processor to service the NTDS interface if necessary. The data are clocked out of this FIFO at 20.48 kHz and converted to analogue. The converted data are filtered with a low pass filter with a –3 dB rolloff located at 8.6 kHz. The full scale output voltage from the D/A stage is 1.0 Volt across an external load of 600 Ohms.

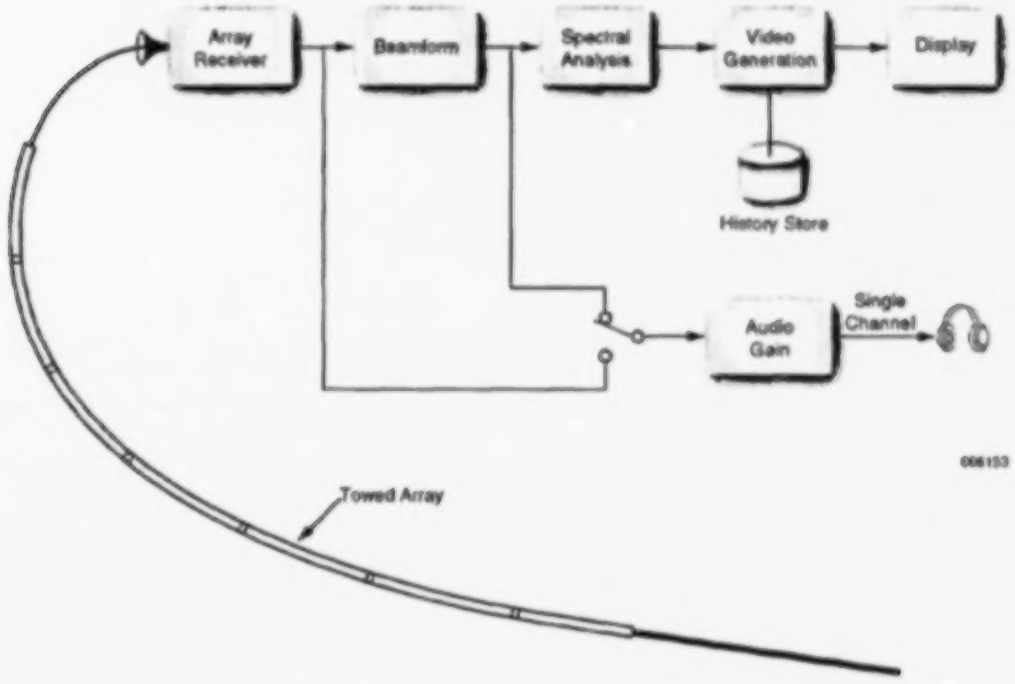


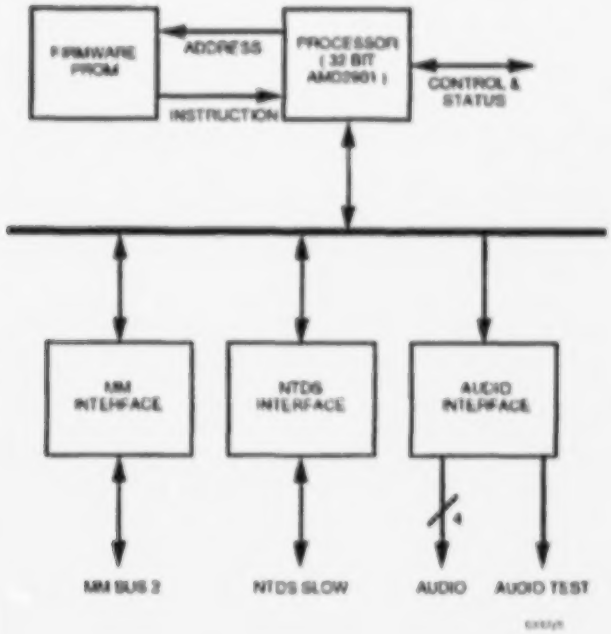
Figure 2. CANTASS a Passive Sonar Block Diagram

3.2.2 *Hardware Limitations*

The deployed or *status quo* CANTASS systems have limited spare processing and a architecture that is difficult to enhance. The CANTASS Signal Processor is loaded at greater than 65%. The audio supports two diotic channels. The firmware used in the Audio CCA (Circuit Card Assembly), see Figure 3, limits the frequency of the audio to less than 1 kHz of 8 bit samples. The operators do not use audio much on the towed arrays at this time. This has been attributed to:

a. **Operator fatigue.** The audio stream sounds like a noise stream with very faint signals buried in this noise. Most if not all operators will tire from listening to this noise. This problem is magnified by the lack of audio recall, see next bullet;

b. **No audio recall.** The operator locates a feature of interest on the video display, i.e., makes the initial detection based on the video presentation. The audio beam being monitored at any given time is not likely pointing in the direction of a new detection (a chance of 1:43). The lack of an audio recall capability removes the audio confirmation capability to help the operator classify or identify the type of new contact. This deficiency is of particular importance if a transient has been detected;



PROPRIETARY
General Dynamics Canada
"Use, duplication or disclosure of this document or any of the information contained herein is subject to the restriction on the title page of this document."

Figure 3. Block Diagram of the Audio CCA (circuit card assembly) used in CANTASS and AN/SQS-510

c. **An inappropriate frequency band.** The audio was restricted to be less than 1 kHz after beamforming, bandshifting etc. This was done to manage data rates within the sonar; and

d. **Limited dynamic range.** The hardware converts 8 bit wide data into the analogue format. More common audio sources, e.g., music Compact Disks (CDs) use a wider 16 bit format to provide a dynamic range better matched to the human ear.

An upgrade project for CANTASS is underway, called CANTASS FORCE. One of the outcomes FORCE is that the towed array data are extracted from the old system and imported into commercial electronics (COTS). This opens the door to new commercial based hardware. The FORCE system design supports expansion by adding extra processing hardware. The configuration of FORCE uses Intel and SUN processors. This upgrade project has been underway for over a year.

3.3 *Hull Mounted Active Sonar*

The active sonar deployed on the frigates and destroyers of the Canadian navy is the AN/SQS-510. This sonar has both active and passive processing capabilities. The two data streams come from the same type of transducer mounted below the keel in a Hull Mounted Sonar (HMS) and in a separately towed body, for a Variable Depth Sonar (VDS). The single operator is seated in front of a console with two colour television monitors, a keyboard with trackball and joystick. The operator is provided with a diotic audio feed that is either from the active processing or the passive stream, see Figure 4.

a. **Passive Band**. The frequency aperture covered by the passive band depends on whether the sonar is configured to ping or not. If the sonar is in passive, the full band[14] (2.2 to 8.6 kHz) is available for processing. If the sonar is in active, i.e., can ping, then the passive band is reduced to the lower portion of the frequency aperture, 2.2 to 5.5 kHz. The passive audio data are only beamformed and filtered to prevent aliasing prior to presentation to the operator. The operator uses the trackball cursor position to define the 'listen direction'. The sonar responds to a direction change within two processing cycles (~ 2 seconds). This slow response can reduce operator effectiveness.

b. **Active Band**. The active band audio is processed by beamforming, bandshifting and adjusted with an automatic gain control. The centre frequency of the echoes is sifted to about 1 kHz. This shift is done to reduce the 8 kHz ping frequency down to a frequency more easily interpreted by the operator. It also increases the percentage frequency shift caused by the Doppler effect for opening or closing targets. Operators can deduce 'up Doppler' or

---

14 In reality the full band is composed of a lower band (2.2 to 5.5 kHz) and an upper band (5.25 to 8.6 kHz). The lower band is produced only in active mode. The processing of the upper band only is performed if pings are not being generated. This prevents pings from contaminating the passive displays and passive audio.

PROPRIETARY
General Dynamics Canada
"Use, duplication or disclosure of this document or any of the information contained herein is subject to the restriction on the title page of this document."

'down Doppler' at this frequency if the sonar is operating in CW. The situation is much more difficult in FM. It is very difficult to determine the Doppler shift from listening to the ping because of the broadband nature of the reverberation. An example of a closing (up Doppler) echo and a low Doppler target can be heard in the following sound clip. The data have been band shifted down to about 1 kHz. The clip is an unusually good sample with higher signal to noise than is commonly found.

Sound Clip: CW ping with Up Doppler and Low Doppler echos

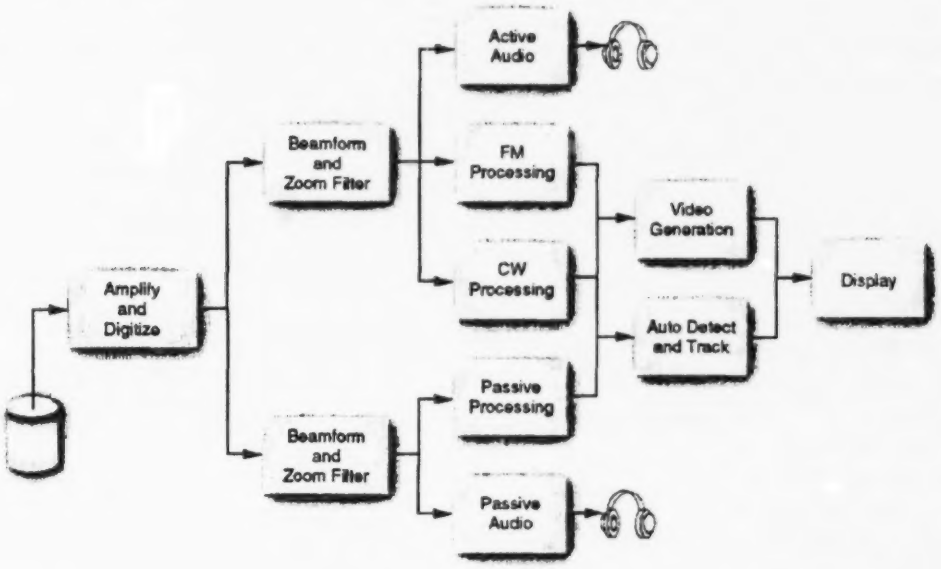


Figure 4. AN/SQS-510 an Active Sonar Block Diagram

3.3.1 Hardware Limitations

Many of the same issues that were identified in the discussion about CANTASS apply to the AN/SQS-510 because of the common hardware. The signal processing is essentially fully loaded and not able to accommodate any increase in processing. The audio channels can not be configured to process for spatial audio. The frequency range of the audio is 8 bits, less than 8.6 kHz.

PROPRIETARY
General Dynamics Canada
"Use, duplication or disclosure of this document or any of the information contained herein is subject to the restriction on the title page of this document."

4. RECOMMENDATIONS

A variety of audio techniques used in sonar have been developed over the years. An emphasis on visual displays for sonar data has reduced most of the traditional uses of audio. Spatial audio reproduction is becoming more common in a wide range of applications: including teleconferencing, simulation and virtual reality (especially PC Games), professional audio and computer music, and advanced machine interfaces for data representation. The advantages of dichotic spatial audio over simple diotic reproduction are significant, e.g., spatial audio improves the legibility and naturalness of a virtual scene, and enables exploitation of spatial auditory cues in order to segregate sounds virtually emanating from different directions. The improvements are expected to be applicable to sonar. Some work has demonstrated the improvements, e.g., reference 6.3g –

> *"The 3 D auditory display was directional and effective in re-creating an acoustical underwater environment. ... Dynamic auditory cues from head movement and visual cues would decrease or eliminate the confusion caused by reversals and inheadedness."*

A number of factors that are outside the domain of spatial audio should be considered as part of a review of audio in general. These factors, if addressed prior to spatial audio implementation, are expected to contribute to an even better audio experience

4.1 Non-Spatial Audio Factors

4.1.1 Improvements to Early Stages of Audio Processing

A number of factors also should be addressed prior to implementing spatial audio. Addressing these issues would improve the effectiveness of the audio channel.

a. The reality is that existing sonars already can generate enough data to completely overwhelm an operator. Future sonars that are being designed have the potential to increase the number of beams by at least a factor of five. This situation demands that some kind of computer based monitor that looks for and alerts the operator of events that should be presented to the operator. Automatic classification is a difficult problem that in general is not solved. All of these factors are made more complex if the signal of interest is a transient. Work is recommended to develop automatic alert capabilities to alert the operator to events of significance.

b. The human ear has a limited bandwidth. In general sonars are capable of providing frequencies that are outside the capability of normal hearing, in particular receiving signals below the comfortable bandwidth of human hearing. Two approaches come to mind to address this reality:

    i. These signals could be band shifted into the frequencies more easily detected. This would preserve the time relationship of the various components – time is not compressed or expanded.

PROPRIETARY
General Dynamics Canada
"Use, duplication or disclosure of this document or any of the information contained herein is subject to the restriction on the title page of this document."

ii. Another way to increase the frequency of the signal is to perform time compression of the signals (similar to listening to tape on fast forward). This fast time replay may provide an appropriate emphasis for certain types of signals.

It is unclear which transformation is the most appropriate. A study is recommended to determine the most appropriate approach.

4.1.2 Operator Fatigue and Historical Data Replay

Audio noise induced fatigue is well known. The operator would be listening to sounds that are mostly noise. In deep water targets are far apart and few in number. This implies long periods of time without contacts – yet the audio would need to be monitored somehow. In shallow water the situation is far more complex. The sonar picture is confused by many possible targets, most of which are not of interest, e.g., wrecks, piers, commercial shipping, geographical features – shoals, sea mounts etc. The sonar operators are already very busy examining the data produced by the sonar.

The recommendation is to reduce the operator stress or fatigue level by allowing the operator to replay audio from the recent past. Should an acoustic event occur, detected either on a video display or in an audio stream, the operator could go back in time and replay the appropriate audio segment to confirm the event's characteristics. This allows the operator to relax a little because of access to an 'instant replay' of recent events. The level of concentration required to manage everything in real time is reduced.

4.2 Spatial Audio Benefits

In the most abstract view, the benefits of spatial audio come when the situation requires the presentation of many sound streams to the operator. What has been done in the past is to add the signals together in a mixer, OMNI-Audio, and let the operator try to figure out the result – a situation that just does not work very well. If spatial audio is used, then the signals are multiplexed into the audio stream in a way the human brain understands and can decode. The operator has a natural way of demultiplexing the data streams back to individual streams. The recovery of the different data streams supports the following benefits:

a. Increased ability to monitor several different acoustic streams of sonar than if an OMNI-audio approach was used. The approach supports the interpretation of data streams that have lower signal to noise ratios (SNR). An upper limit in the SNR improvement is the Array Gain. A typical value for Array Gain is 13 dB for a line array 10 wavelengths long (see footnote 8). This is an upper limit because the operator does not recreate the correct position of the beams as well as the sonar processing.

b. Better separation of voice and acoustic streams. This is superior to piping voice in one ear and sonar acoustic data into the other. Also several different channels can be placed around this virtual audio space – for example: an Active sonar operator to the right, passive sonar operator to the left and supervision above. This leads to a more natural way of interpreting information. The exact perceived position is not too important, just that the various sound streams are separated.

PROPRIETARY
General Dynamics Canada
"Use, duplication or disclosure of this document or any of the information contained herein is subject to the restriction on the title page of this document."

c. The transient detection requires maximising the signal to noise ratio (SNR). But if several channels are to be monitored the traditional method sacrifices the SNR to provide coverage over many channels. The spatial audio allows the monitoring of a number of channels without completely removing the beamforming gain, or array gain, in SNR. This is expected to lead to earlier detection based on transients. A further consequence of the improved SNR is that the characteristics of the transient are more easily interpreted. In other words, the better SNR helps the operator correctly classify the event.

Cost effective processing for producing and manipulating auditory spatial cues, which is required by spatial audio, is available now. This augmented reality must be carefully 'constructed' so that the intuitive understanding of the situation is reinforced not destroyed. For example, headsets, and head tracking improves the localisation of the sounds. The two types of sonar, active and passive, are expected to require different processing approaches to exploit spatial audio technology with the goal to improve the effectiveness of the sonar operator. The various recommendations and issues that require further study are grouped and discussed below.

4.3 Operator or Human Factors

4.3.1 General HRTF or Personal HRTF

Two different types of HRTF are available, generalised or personal, and can be used to generate the spatial audio signals. A generalised HRTF attempts to capture the modification of the sound caused by a generic human head. A personal HRTF is based on detailed measurements made on that specific individual.

Most people perceive separated sound positions when using a generalised HRTF. This is highly significant given the cost and effort required to create a personalised HRTF. A disadvantage of a generalised HRTF is that the accuracy of the perceived sound position is generally poorer compared to a personal HRTF.

On the other hand, the virtual environment created with a personal HRTF comes at a significant cost. Personal HRTFs are derived from binaural measurements inside the ears of the individual seated in an anechoic chamber. This requires a substantial investment in infrastructure and equipment, and is judged at present to be impractical in most applications. Each and every possible sonar operator would need to be measured and the sonar would need to be able to 'upload' the HRTF and then use the specific operator's HRTF.

The recommendation is to use a generalised HRTF and constrain the design of the audio portion of the operator interface to not depend on the detailed placement accuracy of the perceived sonar position.

PROPRIETARY
General Dynamics Canada
"Use, duplication or disclosure of this document or any of the information contained herein is subject to the restriction on the title page of this document."

4.3.2 *Number of Sound Sources*

The number of distinct audio streams multiplexed and presented to the operator needs to be constrained. Multiple voice streams and sonar data streams could easily overwhelm the operator. For example, the active hull mounted sonar can produce 72 active beams and 36 passive beams. One could consider presenting all 108 of these beams to the operator. The resulting audio stream would be so complex that the operator is expected to be overwhelmed.

This is contrasted with the situation in the video displays. The operator has access to representations of all the sonar beams in the form of video display data. The video presentation does not immediately overwhelm the operator because the operator controls the region of the video display being examined, i.e., the operator can zoom in on a specific feature on the video display. Detailed views are available on different video displays as well as summary video displays. The ability to overload the operator with video data still exists, particularly in littoral waters. This overload possibility provides motivation for improving automatic trackers, monitors and alerts in the sonars.

The recommendation is to define an operator interface that allows the operator to define bearings of interest that are to be added to the audio stream. The voice channels would also be controlled by the operator. Research is required to determine the practical limits of numbers and placement of sources. For purposes of discussion, consider the audio being grouped into two stream types – administrative and tactical. The administrative would include general platform messages that have little to do with the immediate tactical situation but are of interest to the operator, e.g., time for dinner. The tactical messages provide a method to exchange information about the sonar situation. The proposal is to place:

a. The administrative messages behind the operator above the horizontal;

b. Tactical messages would be placed to the front also above the horizontal;

c. Operator selected active sonar beams would be placed generally forward on the horizontal plane; and

d. Operator requested passive sonar beams would be placed below the active beams.

Research is required to validate the performance of an operator with this type of sound placement. Variations of this scheme driven by the results of human factor trials would likely lead to adjustments.

PROPRIETARY
General Dynamics Canada
"Use, duplication or disclosure of this document or any of the information contained herein is subject to the restriction on the title page of this document."

4.3.3 Operator Comfort

A number of issues can contribute to operator comfort. Most are well managed with existing headsets. One issue of concern that can be investigated is the possibility of increasing the severity of motion sickness suffered by an operator. This might occur if the realism of the spatial sound is able to conflict with the operator's inner ear motion and orientation sense organs. The recommendation is to perform a study that examines the 'motion sickness' effects of spatial audio on moving platforms.

Another factor, of particular interest in aircraft, is the background ambient noise level. The operator is expected to be able to detect weak signals in an sonar background. This requires that the aircraft generated noise be dramatically reduced. The recommendation is to explore active noise cancellation type of headsets.

4.3.4 Merging of Voice and Data

The operator will always need to have the voice channel available. The addition of audio from a sonar requires care so that the levels are balanced appropriately. The relative audio levels between the various audio streams likely will require adjustment. The recommendation is to perform studies/trials to establish the appropriate levels – both absolute and relative – so the operator effectiveness is enhanced. In some situations, the operator has two headsets - a sonar headset and a voice communication headset. The headsets are typically wired with both ear pieces receiving the same signal, a diotic configuration. The sonar operators should have a single 'stereo' headset that is used for all audio sources, e.g., voice communications, alerts, and sonar data.

4.3.5 Audio Control interface

In general, the operator needs controls that manage the various sound sources. The amplitude of individual audio channels and placement positions of the channels in the virtual space are candidates for operator control. The recommendation is that studies be performed to establish the detailed requirements for the audio control interface and the visual cues that the sonar might provide to re-enforce the otherwise spatial audio experience. The interface control for each type of sonar likely will be unique because of the differences highlighted earlier in this document.

The possibility exists that the various controls suggested above may be able to mitigate the negative attributes of using a generic HRTF by supporting limited customising of the processing. The design requirements of the audio control interface will likely require significant effort but will influence the acceptance of the spatial audio concept.

PROPRIETARY
General Dynamics Canada
"Use, duplication or disclosure of this document or any of the information contained herein is subject to the restriction on the title page of this document."

4.4 Scope of Change

A number of trials or studies have been recommended to help determine the benefits and the parameters that specify the configuration. The studies are expected to take place in a laboratory environment. The results of these studies are required before any modification is attempted to deployed sonars.

The next step would be to introduce a spatial audio demonstration to a deployed sonar. The CANTASS FORCE sonar is recommended as the platform to be used as a demonstrator. This recommendation is believed to manage cost, schedule and impact on the various platforms. CANTASS FORCE is an enhanced version of CANTASS that extracts the raw sensor data from the proprietary hardware and processes the data in commercial platforms (UNIX operating system on SUN platforms or Windows NT on WINTEL platforms). The architecture of the CANTASS FORCE configuration would support the introduction of a spatial audio processor with minimal impact on the existing functionality and existing hardware. The detailed cost and software estimate would require further analysis – well beyond the scope of this contract. This approach would allow the trial configuration to be established with only software changes and new hardware plugged into existing hardware. The new hardware would access the required data over a standard etherNet connection. The new hardware would provide the spatial audio signals for a new headset capable of dichotic operation. Cables would need to be run between the Aft Sonar Instrument Space and the CANTASS Operator positions on the Halifax Class Frigates.

Following a successful demonstration on the CANTASS FORCE configuration the introduction of spatial audio to other sonars is proposed.

PROPRIETARY
General Dynamics Canada
"Use, duplication or disclosure of this document or any of the information contained herein is subject to the restriction on the title page of this document."

5. CONCLUSION

This review has explored issues associated with the presentation of audio to a sonar operator. This review suggests that an opportunity is available that can improve the performance of sonars (including the important Naval acoustic operators).

The *status quo* presents audio in a way that likely will increase operator stress and confusion. This problem is expected to get worse as deployment in littoral waters increases – an environment with a large number of acoustic sources and echoes.

Signal processing (band shifting and/or time compression), data management (instant replay of audio clips) and operator controls (selecting beams of interest and balancing the relative intensities of different audio sources) are proposed to help by increasing operator comfort and reducing operator fatigue. Another type of processing suggested to improve the utility of audio is active background noise cancellation, especially in aircraft.

The use of spatial audio is expected to improve the sonar operator's performance. The importance of spatial audio technology is expected to increase as the quantity of data available to the operator increases. The advantages expected of spatial audio technology include:

a. The operator can monitor for an audio event even if a different type of video data set is being explored, i.e., can always hear regardless of where you're looking.

b. The human ear is superior at detecting transients. The short integration period of the ear is a real advantage for transients – this is contrasted with continuous emissions that benefit from a longer integration period than the ear can provide.

c. Spatial audio multiplexes different audio channels into two signals – one for the left ear and the other for the right ear. The wonderful aspect of this multiplexing is that the operator is able to de-multiplex the signals and restore the separation of the sound. Thus the operator can use a natural ability to focus on sounds from a specific direction in the virtual environment. This spatial filtering is similar to the ability of an individual to focus on a single conversation from a specific individual in a cocktail party environment. This cocktail party effect is real and would be exploited by the sonar.

d. The virtual environment created by spatial audio will have an increased SNR compared to a situation where all the audio channels are just added together. The improved SNR is expected to increase detection ranges, and improve classification of signals.

PROPRIETARY
General Dynamics Canada
"Use, duplication or disclosure of this document or any of the information contained herein is subject to the restriction on the title page of this document."

The proposed work in spatial audio does not remove the need for video displays of sonar data. These displays help the operator overcome weaknesses in the audio presentations such as:

a. Operator fatigue caused by the operator listening to a noise field for long times in search of events of interest.

b. Normal integration periods of the human ear are short compared to integration periods used to detect non-transient emissions from platforms of interest. Video displays and signal processing can create more or less arbitrarily long integration periods.

c. The operator can manage a modest number of acoustic channels, but not as many as most sonar can produce and needs to make decisions about what data should be monitored thereby potentially ignoring vital information. The video displays capture the events from all beams and provide a suitable display and history. The operator need not be looking at the time of the event, because the evidence of an event remains on the screen. If the operator is not listening to the beam – the event would be lost or missed.

In most situations the combination of audio and video information being presented to the operator usually improves the overall understanding of the situation – situation awareness. Historically significant effort has been expended and continues to be expended to design appropriate video display formats. The proposal is to enhance the audio environment to a level similar to the video. The human ear's natural ability to understand transients should to be exploited.

Some of the issues identified in this report require investigation prior to deployment. The general plan is to:

a. Demonstrate the improvement in a laboratory setting prior to a deployed environment.

b. Investigate issues associated with the signal processing and automatic target identification. This work will likely be a requirement for the success of deploying spatial audio in sonars because the operator will need to be alerted to the possibility of an event of interest in a specific beam

c. Create a detailed cost estimate for upgrading equipment to accommodate 3-D audio presentation. This cost may be significant depending on the needs identified in the studies proposed.

PROPRIETARY
General Dynamics Canada
"Use, duplication or disclosure of this document or any of the information contained herein is subject to the restriction on the title page of this document."

6. APPLICABLE DOCUMENTS

The following documents are referenced within the text of this report.

6.1 Canadian Government Documents

Available from:

DMSS - 7
National Defence Headquarters,
555 Blvd de la Carriere
Hull, QC
Canada, J8Y 6R5

a.	C-69-748-000/MB-001	CFTO Operators Manual for the CANTASS Shipboard Electronics Subsystem AN/SQR-502
b.	C-69-799-000/MB-001	CFTO Operating Instructions Detecting-Ranging Set, Sonar, AN/SQS-510 (V) 4

6.2 General Dynamics Canada Documents delivered to the Crown

Available from:

DMSS - 7
National Defence Headquarters,
555 Blvd de la Carriere
Hull, QC
Canada, J8Y 6R5

a.	957249	System Requirement Specification, Shipboard Electronics Sub-System, CANTASS (CONFIDENTIAL)
b.	970552	System Specification of the AN/SQS-510 Sonar

PROPRIETARY
General Dynamics Canada
"Use, duplication or disclosure of this document or any of the information contained herein is subject to the restriction on the title page of this document."

6.3 Commercial Publications

a.	*Principals Of Underwater Sound*	Robert J. Urick, 3rd Edition, 1983, McGraw-Hill, Inc. ISBN 0-07-066087-5
b.	*Underwater Acoustic System Analysis*	William S. Burdic. 1984, Prentice-Hall, Inc. ISBN 0-13-936716-0
c.	*Range discrimination by big brown bats (Eptesicus fuscus) using altered model echoes: Implications for signal processing*	W. Mitchell Masters and K. A. S. Raver J. Acoust. Soc. Am. 107 (1), January 2000 page 625
d.	*Review of Undersea Acoustics to 1950*	Marvin Lasky, J. Acoustic Soc. Am., Vol. 61, No. 2, Feb 1977
e.	*Boundary element method calculation of individual head-related transfer function. I. Rigid model calculation*	Brian F. G. Katz J. Acoustic. Soc. Am, Vol. 110 No.5, Nov. 2001page 2440
f.	*Direct Comparison of the Impact of Head Tracking, Reverberation, and Individualized Head-Related Transfer Functions on the Spatial Perception of a Virtual Speech Source*	D.R Begault, E.M. Wenzel and M.R. Anderson J. Audio Eng. Soc., Vol 49, No. 10, 2001 October page 904
g.	*3_D Auditory Display of Passive Sonar Data*	S Richardson and C Loeffler Proceedings of 137th Meeting of the Acoustic Soc. of Am and the 2nd Convention of the European Acoustics Assoc.: Forum Acusticam Berlin, March 14-19, 1999. ISBN 3-9804568-5-4
h.	*LCD Blink Effect: Acoustic specific or widespread phenomenon?*	G. Wright, F. Bosko, S Allen, and T Thomas Proceedings of SPIE Cockpit Displays VII: Displays for Defense Applications, Volume 4022, August 2000. (the Abstract of this paper is available at: http://spie.org/scripts/toc.pl?volume=4022&journal=SPIE.&qs=spie&ab

PROPRIETARY
General Dynamics Canada
"Use, duplication or disclosure of this document or any of the information contained herein is subject to the restriction on the title page of this document."

DOCUMENT CONTROL DATA SHEET

1a. PERFORMING AGENCY	2. SECURITY CLASSIFICATION
General Dynamics Canada, 3785 Richmond Rd., Nepean ON, K2H 5B7	UNCLASSIFIED Unlimited distribution -
1b. PUBLISHING AGENCY DCIEM	

3. TITLE

(U) Uses of spatial audio in sonar

4. AUTHORS

Brian E. Cooke

5. DATE OF PUBLICATION	6. NO. OF PAGES
February 28 , 2002	29

7. DESCRIPTIVE NOTES

8. SPONSORING/MONITORING/CONTRACTING/TASKING AGENCY
Sponsoring Agency:
Monitoring Agency:
Contracting Agency : DCIEM
Tasking Agency:

9. ORIGINATORS DOCUMENT NO.	10. CONTRACT GRANT AND/OR PROJECT NO.	11. OTHER DOCUMENT NOS.
Contract Report CR 2002-054	1CV	Contract No.: W7711-017719/001/TOR

12. DOCUMENT RELEASABILITY

Unlimited distribution

13. DOCUMENT ANNOUNCEMENT

Unlimited announcement

14. ABSTRACT

(U) Underwater submarine warfare has imposed an increased threat in present-day military operations. The Canadian Department of National Defence and the defence forces in allied nations are presently investigating commercial off-the-shelf technology for enhancing the presentation of sonar data in the auditory modality in order to improve operator efficiency. One identified technology is three-dimensional (3-D) audio. A 3-D sound gives the listener the perception of the sound emanating outside of his/her head at a particular point in space even though the signal is presented over headphones. The use of 3-D audio as opposed to conventional aural presentation of sonar data is expected to result in improved operator detection and classification of enemy targets. Some recommendations and their implications are discussed in this report.

(U) La guerre sous-marine mettant en jeu des sous-marins représente une plus grande menace pour les opérations militaires d'aujourd'hui. Le ministère canadien de la Défense nationale et les forces de défense des nations interalliées étudient actuellement la technologie de série servant à améliorer la présentation des données sonar dans la modalité auditive afin de rendre l'opérateur plus efficace. Une des technologies identifiées est l'audio tridimensionnel (3-D). Un signal sonore 3-D donne à la personne qui écoute l'impression d'entendre un son émanant d'une source hors de sa tête, situé à un point particulier de l'espace, bien que le signal passe par des écouteurs. On espère que l'utilisation de l'audio 3-D au lieu de la présentation sonore classique des données sonar réussira à améliorer la détection et la classification meilleures des objectifs ennemis par l'opérateur. Certaines recommandations et leurs implications font l'objet du présent rapport.

15. KEYWORDS, DESCRIPTORS or IDENTIFIERS

(U) sonar; three-dimensional audio; auditory displays